

·综 述·

人工智能在超声医学领域中的应用

刘睿峰，夏 宇，姜玉新

中国医学科学院 北京协和医学院 北京协和医院超声医学科，北京 100730

通信作者：夏 宇 电话：010-69159311，E-mail：yuxiapumch@ aliyun.com

【摘要】人工智能（artificial intelligence，AI）近几年再度成为各领域关注的焦点，其中深度学习的提出带来了一系列革命性变化，而随着计算机视觉向深度学习过渡以及硬件和大数据的进步，AI 在图像识别领域展现出更广阔的发展前景。深度学习模型使得相关图像算法甚至达到了比人眼更高的识别准确率，这为医学影像的发展提供了巨大契机。超声医学作为影像领域的重要分支，利用 AI 相关算法进行声像图分析的研究不断涌现，不仅为临床科研提供了新思路，亦有助于提高超声诊断的准确性。

【关键词】人工智能；深度学习；超声成像；临床研究

【中图分类号】R445. 1 【文献标志码】A

Application of Artificial Intelligence in Ultrasound Medicine

LIU Rui-feng，XIA Yu，JIANG Yu-xin

Department of Ultrasound，Peking Union Medical College Hospital，Chinese Academy of Medical Sciences & Peking Union Medical College，Beijing 100730，China

Corresponding author：XIA Yu Tel：010-69159311，E-mail：yuxiapumch@ aliyun.com

【**Abstract**】 As deep learning brings a series of revolutionary change，artificial intelligence has become a focus in various fields again in recent years. With the transition from computer vision to deep learning and the progress in hardware and big data，artificial intelligence，has demonstrated broader prospects for the development of image recognition. Image algorithm exploiting deep learning model has achieved better identification accuracy than the naked eye，which offers the possibility of making breakthrough in medical imaging field. Ultrasonography is a main branch of medical imaging. An increasing number of papers on research of the application of artificial intelligence-related algorithms into analyzing ultrasonographic images provide new insights into clinical research. Meanwhile，specific software is able to compensate for the practitioner's deficiency in experience and improve diagnostic accuracy as well.

【**Key words**】artificial intelligence；deep learning；ultrasonography；clinical research

1950 年 Alan Turing 发表的《计算机器与智能》一文中提出了“图灵测试”，首次预见性地展示机器与人类难以分辨的智能行为[1]；6 年之后，达特茅斯大学研讨会上科学家们对麦卡锡提出的新术语“人工智能（artificial intelligence，AI）”初步认同并接受，标志着这一概念的正式诞生。随后几十年间，AI 经历了从跳棋程序到专家系统，从发展逻辑推理的第 5 代计算机研制计划乃至今天基于网络大数据和深度学习模型的分布式 AI 研究等阶段，已渐趋成熟。Stuart Russell 在其 1995 年出版的著作《人工智能：一种现代方法》中将 AI 定义为有关“智能主体研究与设计”的学问，而智能主体是指可以观察周遭环境并作出行动以达到目标的系统[2]。自 2016 年 3 月 Alphago 与围棋世界冠军李世石的人机大战后，AI 又一次进

入公众视野，备受关注。基于医学影像特点和提高影像诊断效能的诉求，AI 有望在将来成为影像医生诊治过程中的有效辅助工具。本文将结合具体案例介绍 AI 在超声医学领域的可行性和局限性，并对未来应用前景进行分析与预测，希望为进一步研究提供依据。

1　人工智能与医学影像

鉴于医学影像领域信息更加结构化，大部分基于图像的判断相比临床电子病历总结式的描述更加客观，深度学习取得突破前就有学者利用 AI 方法中的人工神经网络来分析医学图像，然而受限于梯度消失和过拟合问题，模型很难构建深层次的架构来学习，同时还要面对数据量和终端计算能力不足的困难，相关研究取得的成果有限。而如今，基于放射信息系统（radiology information system，RIS）、医学图像存档和传输系统（picture archiving and communication system，PACS）中的大量病例图像，AI 可更好地进行深度学习，同时利用深度学习中一系列具有不同优势和适用性的模型，如循环神经网络、卷积神经网络等，建立输入图片特征与输出目标结果之间的对应关系，更加高效识别图片中的对象，为相关研究与发展提供了更多可行性。

深度学习作为此次 AI 兴起以及相关医学影像研究兴盛的核心技术，其主流模型延续了早期机器学习中的一个重要算法，即人工神经网络，原理是模拟人脑多层神经网络结构并作出判断，而应用于医学影像的深度学习以神经网络为主，大致分为两类：（1）监督学习，包括递归神经网络、卷积神经网络和普通深度神经网络等主流模型；（2）无监督学习，包括深度生成模型（预训练）和自编码器等模型。二者主要区别在于对具有相关特征的训练数据进行分类时，训练数据是否有人工标注的标签，如数据有标签则称为监督学习，无标签即为无监督学习，又称聚类。监督学习中的卷积神经网络正是目前医学影像领域 AI 的研究热点，作为一种具有自主学习能力的神经网络，其多层结构可基于多级抽象提取一系列辨识性特征，从而识别图像中的目标，超声医学中的前沿研究方法也多以此为核心。该神经网络模型对于解决现阶段科研中分类、检测、分割这 3 个主要问题均具有重要意义，其中分类问题主要涉及完成探测组织结构异常并将其划归至各个疾病类别中的任务，类似于诊断的思维过程，是最典型的科研问题。

2　人工智能在超声医学领域的应用

2.1　传统人工神经网络研究

2.1.1　大脑中动脉痉挛所致狭窄诊断方法效能评估

大脑中动脉痉挛是蛛网膜下腔出血的严重并发症，早期诊断和干预对预防卒中十分重要。数字减影血管造影（digital subtraction angiography，DSA）是作出该诊断的金标准，但鉴于其为有创性操作并可导致相关并发症，不宜作为常规监测手段。经颅多普勒（transcranial Doppler，TCD）作为一项非侵入性监测方法，可在床旁操作，且指导临床实践准确性高，因此被广泛应用[3]；同时，有文献显示，经颅双功能彩色多普勒超声（transcranial color-coded duplex sonography，TCCS）在大脑中动脉痉挛所致狭窄的诊断上比 TCD 具有更高准确性[4]，但这些文献存在 TCD 和 TCCS 入组人群不同、研究组与对照组基线特征不同、受试者工作特征曲线（receiver operating characteristic curve，ROC）仅包括一项脑动脉血流动力学参数等不足，横向比较存在较大争议，研究结果缺乏说服力。于是，Swiercz 等[5]利用传统人工神经网络建模，并将经 TCD 和 TCCS 获得的数据处理后的输出值与 DSA 的结果进行匹配，以匹配程度最高的模型作为评判 TCD 和 TCCS 诊断准确性的仲裁者，该模型能够将大脑中动脉平均流速（mean velocity，Vmean）、收缩期峰值流速（systolic peak velocity，Vps）及舒张末期流速（end of diastolic velocity，Ved）整合为一个集合参数，建立 ROC 并进行相关比较，从而避免了超声医生解读数据作出判断时的主观影响，减少了混杂因素，但该研究由于样本量有限，阳性病例数相对较少，模型的准确性受到质疑。

2.1.2　建立非侵入性肝脏纤维化评级体系

慢性肝炎或肝损伤所致肝纤维化是肝硬化的共同特征，肝纤维化是可逆的病理过程，及时有效的治疗可避免其进一步发展为肝硬化。为了在治疗过程中定期进行纤维化或硬化评级，常常选择超声作为监测手段。Zhang 等[6]将传统人工神经网络的训练规则调整为错误反向传播算法，即将输入值包括肝脏实质、脾厚度、肝动脉搏动指数、衰减指数及肝静脉频谱经过一系列不可知的调整后与已有的肝纤维化分级进行比较，如有偏差则返回上一步的层级训练，直至单位错误总和被调整至最小，训练出的相应模型经过评估，

其准确性可达 88.3%（曲线下面积 0.9222）。相较肝活检，结合了 AI 技术的超声监测不存在样本取材问题以及医师经验水平差异所导致的误差，为慢性肝病患者的临床决策提供了另一种参考。

2.1.3 早期甲状腺结节良恶性分类

类似的基于传统人工神经网络的研究还涉及甲状腺结节良恶性诊断，由于医疗资源普及和定期体检的推广，许多甲状腺结节被发现，为进一步定义其良恶性，不得不施行细针穿刺活检或定期随访，给医院诊疗和患者均增加了负担，为了提高单次诊断的准确性，Zhu 等[7]利用具有显著统计学意义（$P<0.001$）的 6 项指标，包括形状（纵径长于横径）、边界（界线模糊）、回声（低回声）、质地（实质）、钙化（微钙化）、晕环（缺如）作为输入值训练传统人工神经网络模型以预测是否有恶性结节的存在，其准确性、敏感性、特异性分别可达 82.3%、84.5%和 79.1%。

2.2 以卷积神经网络为主的深度学习模型

2.2.1 处理分类问题

甲状腺影像报告和数据系统（thyroid imaging-reporting and data system，TI-RADS）联合深度学习分辨甲状腺结节良恶性：TI-RADS 现已广泛用于分类甲状腺结节并对其恶性风险进行评分，然而这种方法耗时、费力且通常不够稳健，其诊断的准确性不仅受检查医生个人经验影响，且受结节回声变异性等原因限制。Acharya 等[8]基于离散小波变换特征，利用计算机对不同模态下的声像图进行归类，取得了 98.9%～100%的准确性，这一类新的“计算机决定的特征”完全不同于临床微钙化灶等分类经验，为深度卷积神经网络的应用开创了前路。与传统特征提取方法相比，深度卷积神经网络具有两个优点：(1) 针对操作时不同灯光条件、垂直和水平位移等所致的诸如形状等失真改变，表现更加稳定；(2) 特征提取时耗费更少的计算资源，故结合预处理和微调等方法后，训练出的 AI 在二维声像图上识别甲状腺良恶性的准确性、灵敏性和特异性分别可达 96.34%、82.8% 和 99.3%，明显优于传统方法[9]。

2.2.2 处理检测问题

协助定位从而为急性阑尾炎诊断提供更多证据：急性阑尾炎是外科常见急症，但典型症状发生率仅为 66%，其临床诊治存在较多难点，在儿童、老年人、孕妇等特殊群体中常因诊断困难而致穿孔。超声检查在疑似病例，特别是其他检查不支持但患者又有持续症状的诊断中提供了有价值的参考信息。一系列征象如阑尾外径>6 mm、钙化灶及其壁上血流信号增强、管腔内积液均可为难以确诊又高度怀疑阑尾炎的患者增加诊断依据，其中阑尾外径在探头加压探查下>6 mm作为诊断时准确性最高的超声征象[10]。有研究利用一系列 AI 算法除噪，提高亮度对比，提取出更加完整的筋膜底线以确定阑尾所在区域，并避免了大量腹水影响，其中的核心算法 Fuzzy ART 是一种无监督神经网络学习算法，亦属深度学习模型的一种，具有实时学习、无固定目标值、对经验知识归类等特点，其应用显著提高了确定阑尾区域的准确性，在已确诊阑尾炎的 40 幅声像图中，最新算法可成功识别其中 38 幅，真阳性率可达 95%，与 CT 诊断能力相当，且无辐射[11]。

2.2.3 处理分割问题

劲动脉相关参数及病变测量：颈动脉内-中膜厚度（carotid intima-media thickness，CIMT）对预测心血管病风险十分重要，超声检查时需医师测量声像图中远端血管壁管腔与内膜交界至中膜与外膜交界的距离，此人工标记耗时乏味，于是有研究者提取 92 份 CIMT 检查录像作为数据库，每一份均由专家划出 3 个感兴趣区（region of interest，ROI），之后再在 276 份 ROI 中分别标识出管腔-内膜交界平面和内膜-中膜交界平面作为训练的图像块进而建立卷积神经网络，经过特定训练的卷积神经网络即可自动分割出需要的界面交由电脑测量，结果优于人工测量[12]。也有学者设计调整了卷积神经网络模型用于颈动脉斑块的测量，并且相较以纹理特征作为输入支持向量机（support vector machine，SVM）分类器的传统机器学习方法，卷积神经网络的结果在准确度（0.9733 比 0.9638），敏感性（0.9653 比 0.9746），特异性（0.9720 比 0.9602），马修相关指数（0.9444 比 0.9254），约登指数（0.9466 比 0.9256）方面均更优[13]，提示了卷积神经网络在临床图像识别中应用的可行性与巨大潜力。

2.2.4 卷积神经网络迁移学习

中孕期胎儿畸形筛查标准平面自动定位：产前畸形筛查是胎儿超声检查的重要内容，不同医疗机构敏感性从 27.5%至 96%不等[14]，其中很重要的一个影响因素即为标准平面的获取，通常需要医师具备全面的解剖知识和大量的经验。胎儿腹围平面（fetal abdominal standard plane，FASP）的自动定位因涉及复杂的解剖结构，可获取的样本量较少，缺乏足够的数据来训练卷积神经网络模型，这一直是一个充满挑战性的难题，于是有学者利用已被大量不

相干原始图片训练过的卷积神经网络的较低层级网络作为基础，将其中的经验知识转移到用来识别定位 FASP 的卷积神经网络模型中，其表现明显胜过之前的基于少量数据训练的低维度模型，准确率、精确率、召回率和综合评价指标分别达到了 0.904、0.908、0.995 和 0.950[15]。科研中使用患者的数据常常涉及隐私等伦理问题，需征得患者同意，这就大大减少了可使用的图像素材；且鉴于医学影像图像对比度较低，训练数据量级要求不高，利用由其他图片（通常是自然图片）训练生成的神经网络进行迁移学习来处理这类问题也很常见。这种方法很好地处理了因数据过少导致的模型过拟合问题，将卷积神经网络在临床应用的可行性进一步拓展，展现了其处理医学图像的可预见前景。

2.2.5　超声心动图相关综合应用

涉及二维及多普勒超声心动图的研究常常伴随大量的测量参数，这些参数对于评价心肌及瓣膜的结构和功能十分重要。然而在实际工作中，理清并针对性地使用如此巨大的潜在数据组合对于一个忙碌的临床医生可行性不高，而且操作者的手法、技巧、经验对于最终测量结果影响均较大，甚至由于不同时间操作者本身状态不同，同一个人所测量的结果可重复性也很难得到保证。相关研究将供训练的数据集（心脏四腔心切面图）由心脏专家按不被接收（0 分）至极佳（5 分）进行总体质量评分，利用卷积神经网络模型更不容易过拟合、更容易训练、加权调整更少的特性，根据上述数据集即时生成自动化回声评分（automated echo score，AES）作为质量反馈[16]，促进低年资医生获取更高质量、更标准的相应切面声像图，提高了医疗效率；同时也有研究显示运用算法对左心室射血分数进行自动测量计算，相对于传统的手工裸眼标记测量，平均处理时间仅需 8 s，且可重复性高，结果相对精确[17]；针对此类研究尚存的争议：算法和计算机技术领域的矛盾，医学方面比如金标准、心室肌小梁等问题[18]，有研究者使用融合了机器学习算法的软件对二尖瓣相关的 6 项参数：二尖瓣环三维前外后内径、二尖瓣环三维前后径、瓣环面积、非平面角度、二尖瓣环总周长、总瓣叶进行测量比较，观察检测器本身是否会影响参数，设定 $P=0.0083$，无影响为无效假设，得到 6 项参数的相应 P 值分别为 0.72、0.25、0.07、0.03、0.13、0.15，结果显然瑕不掩瑜，即使测量数据存在微小的不一致性，研究者也相信软件可以通过自主学习适应调整以改善表现[19]。更多的研究也提示未来机械臂与影像自动化分析的共同发展，有望实现无人干预的全自动化超声心动图的获取、识别和定量分析，在可携带计算机上施行实时超声心动图分析其实已有据可循[20]，同时深度学习相关算法模型的嵌入，相较于现阶段研究中传统的其他机器学习方法，有望为 AI 在心脏疾病诊断方面带来更广阔的发展前景。

2.3　超声医学科产品化应用实例

我国自主研发的 DE-超声机器人已投入临床进行试验，这是一款基于超声影像，辅助医生进行甲状腺结节良恶性识别的智能诊断系统。测试环节中选取省级三甲教学医院不同级别大夫对同一帧图像进行识别，诊断准确率在 60%~70%，该设备协助医生工作同步进行诊断，自动采集图像并给出结果及其概率值，准确性可高达 87%。总的来说，该系统大幅减轻了医生工作量，提升了影像诊断的精准性，既可节约医疗资源及社会成本，又可支持国家分级诊疗医改战略，但由于智能程序潜在的不稳定性及运算错误，同时考量到医学的发展与进步，如何安全有效在工作中使用甚至推广该类 AI 产品仍需进一步探索。

3　前景与局限

随着医院管理信息化和智能化水平的不断提高，AI 与医疗行业的结合已是未来医疗发展的必然方向。2017 年 2 月 17 日，国家卫生和计划生育委员委发布了 2017 版“人工智能辅助诊断技术管理规范”及“人工智能辅助诊断技术临床应用质量控制指标”，提示 AI 已经真正开始走入临床工作并引起相关部门重视，但这是一个多学科交叉合作的过程，需要广大高水平专科医生积极参与，提供高质量数据集训练 AI，并为保障相关应用的准确性和稳定性进一步完善评估方案。

AI 在医疗领域的应用已有相对成熟的系统如“沃森（IBM 公司）”作为参考，由于医学影像领域的特殊性，例如数据特征、人文交互少等，AI 的应用阻力相对较小，国内外涉及图像区域分割、图像目标检测、图像配准等多个领域的研究也逾渐火热，相关模型以深度学习为主，结合机器学习多种方法，其发展对科研和临床的助力不可小觑，但由于不同组织变异性大，边界模糊，微细血管神经分布复杂等原因，通常仅针对某一特定器官或疾病，

多种类型方法结合与改进使用以便相互弥补算法缺陷，尚未有普适的方法可以对任意一张超声声像图进行解析[21]。

关于 AI 是否会取代医生的问题，应基于 AI 应用于超声诊断需要完成两项基本任务，即扫查和读片进行分析。目前更多研究集中于读片场景的应用，而针对患者的操作扫查研究，有学者提出可用一个计算机控制的机械臂控制探头的位置和角度进行不同标准切面的扫查，根据高矮胖瘦不一，机械臂借助一些辅助参数和标记部位精确调整扫查图像的位置和范围从而得到标准切面图像[22]。但是实际工作中，一个超声诊断临床决策的实现往往要结合多个切面、直接与间接征象以及临床信息的分析，这样的能力或许需要到强 AI 阶段才能实现。

一种新药从研发到推广应用至少需要经过 3 个阶段，表明医疗创新实例往往需要面临更高的门槛，这主要是由于医疗行业的预防原则所致。AI 亦是如此，其需要经历大量的测试，测试越多，暴露的信息与问题越多，继而越能够预测后续可能产生的风险，同时考虑到法律法规、民众接受度以及潜在的医疗事故背后相应的权责等问题，在大量测试后 AI 能否切实应用于实际工作尚未可知，而如果 AI 能够代替医生完成大部分工作，医生就可将工作重心放在临床决策的审核及科学研究方面，这一应用将解放医生们的生产力，使医生更加具有创新动力，为专科发展带来裨益！

参 考 文 献

[1] Turing AM. Computing machinery and intelligence [J]. Mind, 1950, 59: 433-460.

[2] Stuart JR, Peter N. Artificial intelligence: modern approach [M]. Englewood: Prentice Hall, 1995.

[3] Sekhar L, Wechsler L, Yonas H, et al. Value of transcranial doppler examination in the diagnosis of cerebral vasospasm after subarachnoid hemorrhage [J]. Neurosurgery, 1988, 22: 813-821.

[4] Baumgartner RW, Mattle HP, Schroth G. Assessment of ≥50% and <50% intracranial stenoses by transcranial color-coded duplex sonography [J]. Stroke, 1999, 30: 87-92.

[5] Swiercz M, Swiat M, Pawlak M, et al. Narrowing of the middle cerebral artery: artificial intelligence methods and comparison of transcranial color coded duplex sonography with conventional TCD [J]. Ultrasound Med Biol, 2010, 36: 17-28.

[6] Zhang L, Li QY, Duan YY, et al. Artificial neural network aided non-invasive grading evaluation of hepatic fibrosis by duplex ultrasonography [J]. BMC Med Inform Decis Mak, 2012, 12: 55.

[7] Zhu LC, Ye YL, Luo WH, et al. A model to discriminate malignant from benign thyroid nodules using artificial neural network [J]. PLoS One, 2013: e82211.

[8] Acharya UR, Swapna G, Sree SV, et al. A review on ultrasound-based thyroid cancer tissue characterization and automated classification [J]. Technol Cancer Res Treat, 2014: 289-301.

[9] Chi J, Walia E, Babyn P, et al. Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network [J]. J Digit Imaging, 2017, 30: 477-486.

[10] Kessler N, Cyteval C, Gallix B, et al. Appendicitis: evaluation of sensitivity, specificity, and predictive values of US, doppler US, and laboratory findings [J]. Radiology, 2004, 230: 472-478.

[11] Kim KB, Park HJ, Song DH, et al. Developing an intelligent automatic appendix extraction method from ultrasonography based on fuzzy ART and image processing [J]. Comput Math Methods Med, 2015, 2015: 389057.

[12] Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? [J]. IEEE Trans Med Imaging, 2016, 35: 1299-1312.

[13] 孙夏，吴蔚，吴鹏，等. 基于卷积神经网络的颈动脉斑块超声图像特征识别 [J]. 中国医疗器械信息，2016, 9: 4-8.

[14] Salomon LJ, Winer N, Bernard JP, et al. A score-based method for quality control of fetal images at routine second-trimester ultrasound examination [J]. Prenat Diagn, 2008, 28: 822-827.

[15] Chen H, Ni D, Qin J, et al. Standard plane localization in fetal ultrasound via domain transferred deep neural networks [J]. IEEE J Biomed Health Inform, 2015, 19: 1627-1636.

[16] Abdi AH, Luong C, Tsang T, et al. Automatic quality assessment of echocardiograms using convolutional neural networks: feasibility on the apical four-chamber view [J]. IEEE Trans Med Imaging, 2017, 36: 1221-1230.

[17] Knackstedt C, Bekkers SC, Schummers G, et al. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain: the fast-EFs multicenter study [J]. J Am Coll Cardiol, 2015, 66: 1456-1466.

[18] Furiasse N, Thomas JD. Automated algorithmic software in echocardiography: artificial intelligence? [J]. J Am Coll Cardiol, 2015, 66: 1467-1469.

[19] Jeganathan J, Knio Z, Amador Y, et al. Artificial intelligence in mitral valve analysis [J]. Ann Card Anaesth, 2017, 20: 129-134.
[20] Kumar S, Nilsen WJ, Abernethy A, et al. Mobile health technology evaluation: the health evidence workshop [J]. Am J Prev Med, 2013, 45: 228-236.
[21] 王弈，李传富. 人工智能方法在医学图像处理中的研究新进展 [J]. 中国医学物理学杂志，2013，3：4138-4143.
[22] Priester AM, Natarajan S, Culjat MO. Robotic ultrasound systems in medicine [J]. IEEE Trans Ultrason Ferroelectr Freq Control, 2013, 60: 507-523.

（收稿日期：2017-09-06）